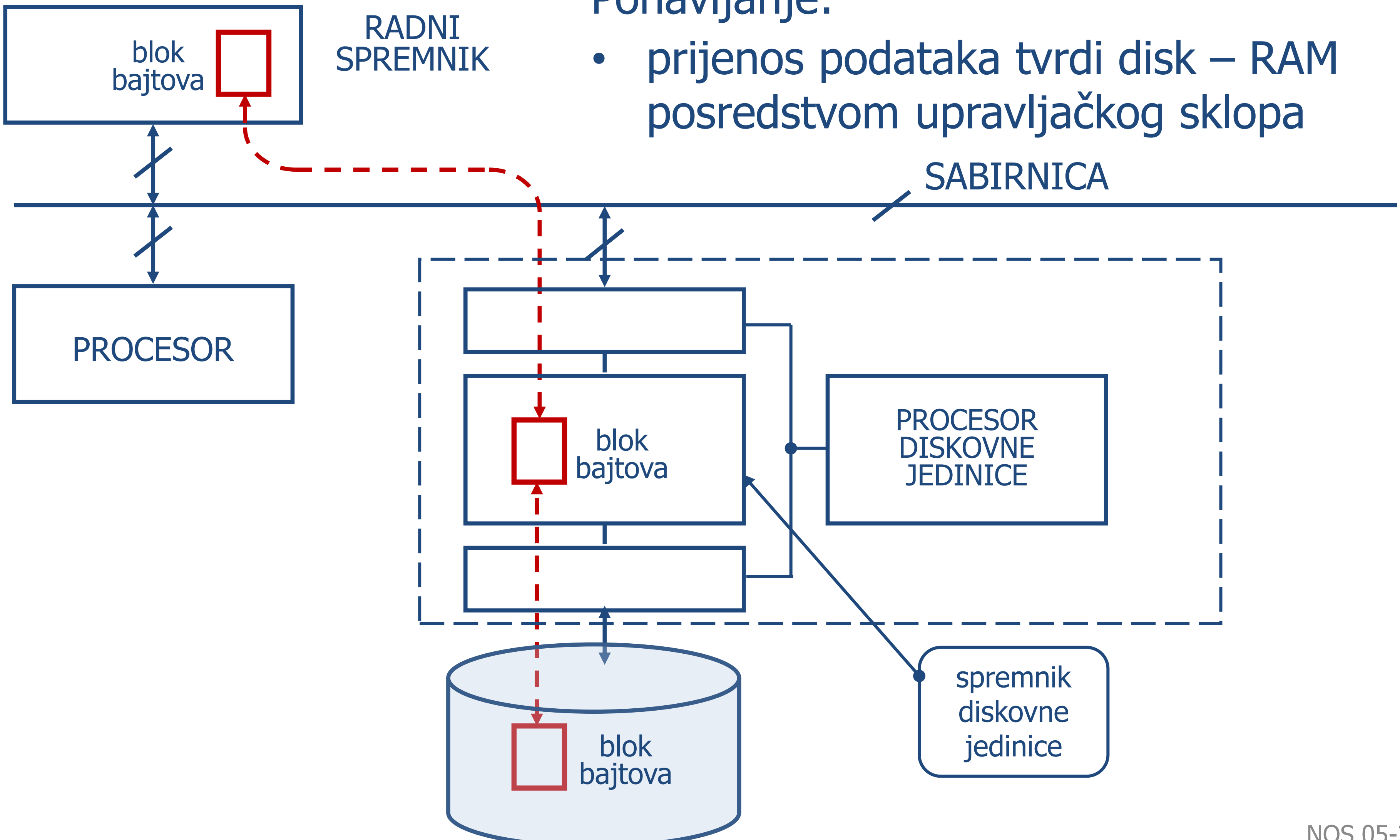
5. Višediskovni zalihosni spremnici

RAID - Redundant Array of Independent (Inexpensive) Disks

Dvanaesto poglavlje u udžbeniku L. Budin, M. Golub, D. Jakobović, L. Jelenković, Operacijski sustavi

Ponavljanje:

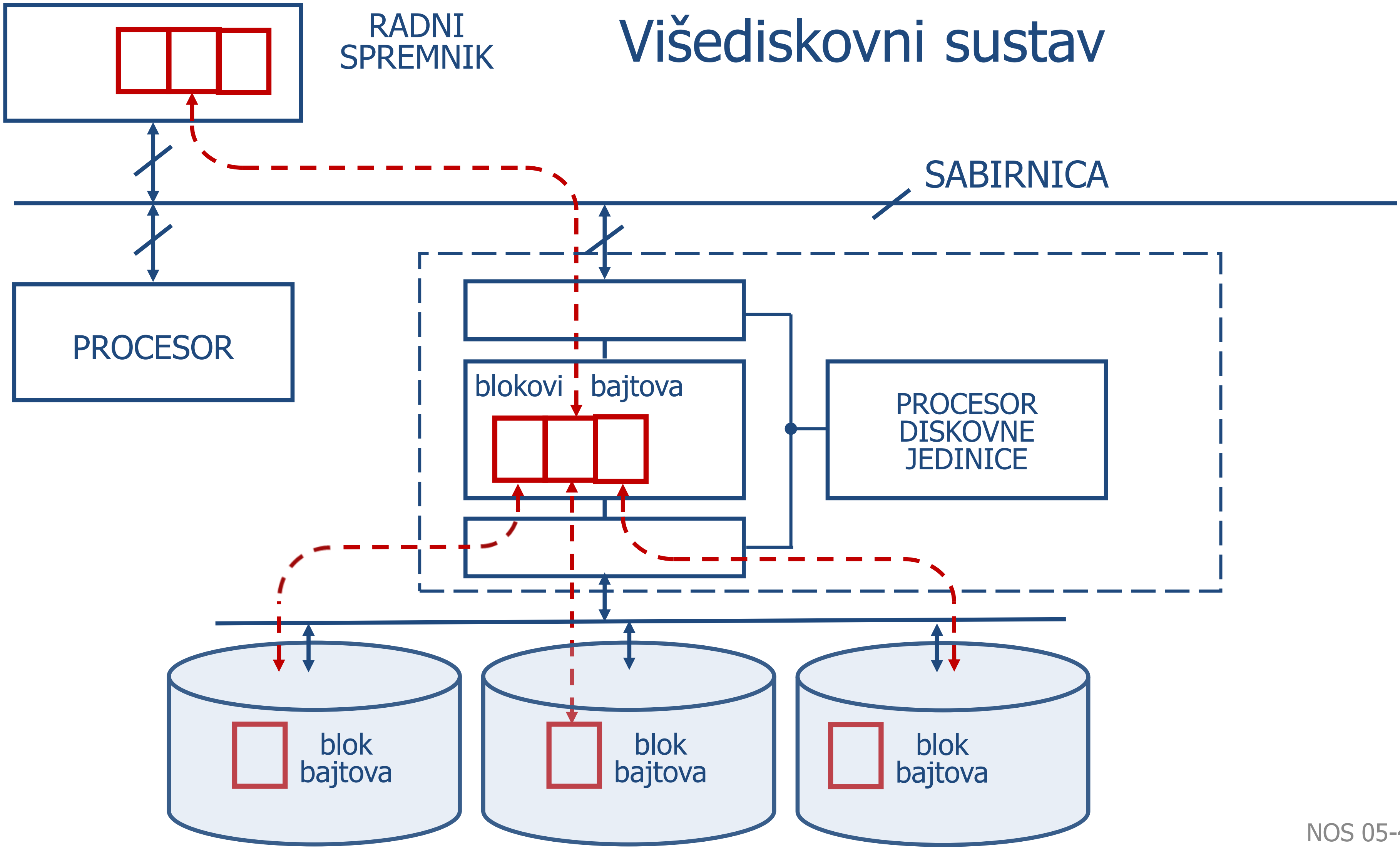
- prijenos podataka tvrdi disk – RAM posredstvom upravljačkog sklopa



U čemu je problem?

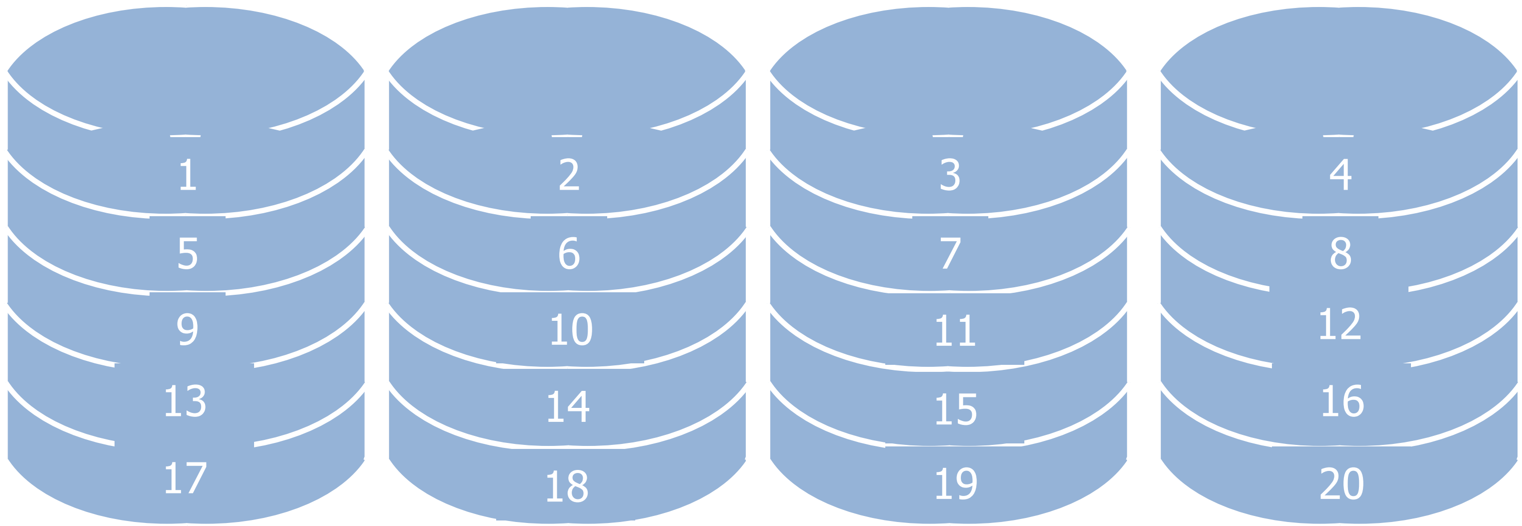
- trajno unapređivanje svojstava procesora i spremnika, a s druge strane, napredak svojstava magnetskih diskova mnogo je sporiji
- defragmentacijom diskova može se ubrzati pristup disku
  - raspršeni smještaj blokova podataka minimizira fragmentaciju diska
- ubrzanje pri prijenosu može se, primjerice, postići tako da se:
  - umjesto prijenosa jednog bloka u spremnik upravljačkog sklopa prenijeti sadržaj cijele staze
  - posluživanje više zahtjeva minimizacijom pomake glava
- Sva ta nastojanja nisu dovoljna!
- Rješenje: smještaj podataka na više diskova kojima se može istodobno pristupati

Višediskovni sustav



Pojasna organizacija višediskovnog adresnog prostora

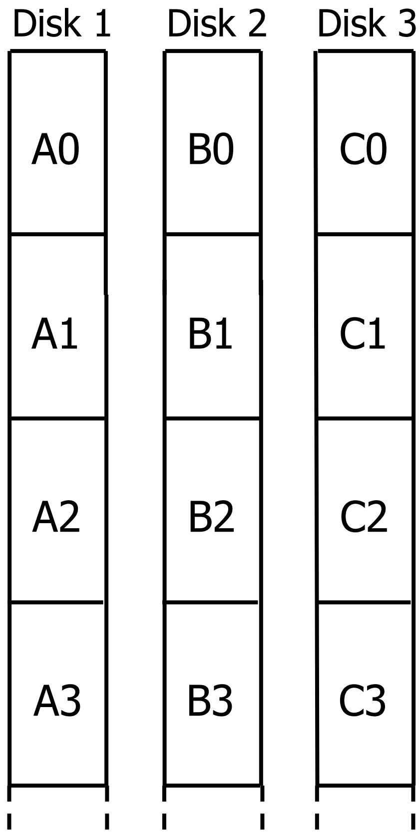
- višediskovni sustav sastoji se od polja diskova (engl. *disk array*) kojima se pristupa paralelno
- polje diskova promatra se kao kao jedan logički adresni prostor
- osnovna jedinica adresiranja podatkovni pojas (engl. *data stripe*)



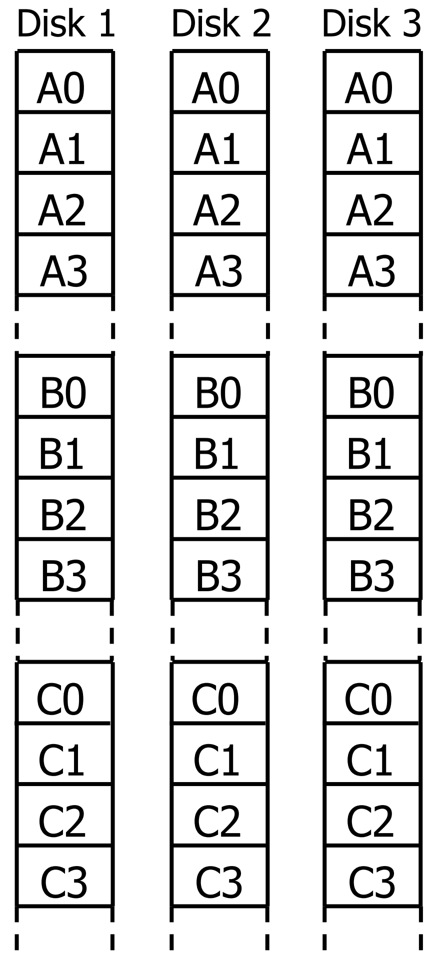
Vrste pojasne organizacije

- sitno zrnata (engl. *fine-grained striping*)
  - podaci se podjednako raspoređuju na sve diskove
  - najmanja veličina pojasa = veličina jednog sektora (u literaturi se spominju bajtovi ili čak bitovi)
  - uvijek se podjednako angažiraju svi diskovi
  - brzina posluživanja zahtjeva je povećana za faktor koji je jednak broju diskova
- krupno zrnata (engl. *coarse-grained striping*)
  - jedan pojas sastoji se od nakupine sektora
  - nije nužno da svi diskovi sudjeluju u ispunjenju svakog zahtjeva
  - polje diskova može istodobno ispunjavati više manjih pojedinačnih zahtjeva
  - prilikom prijenosa velikih količina podataka opet se angažiraju svi diskovi

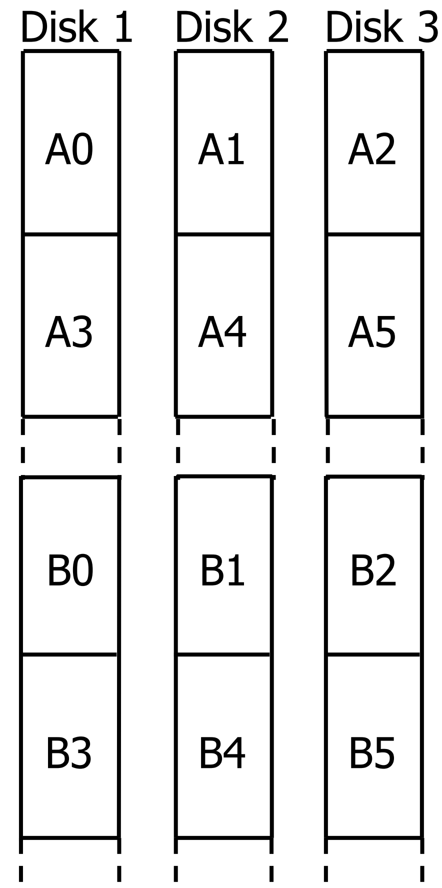
Smještaj podataka na pojedinačne diskove



Sitnozrnati pojasni smještaj



Krupnozrnati pojasni smještaj

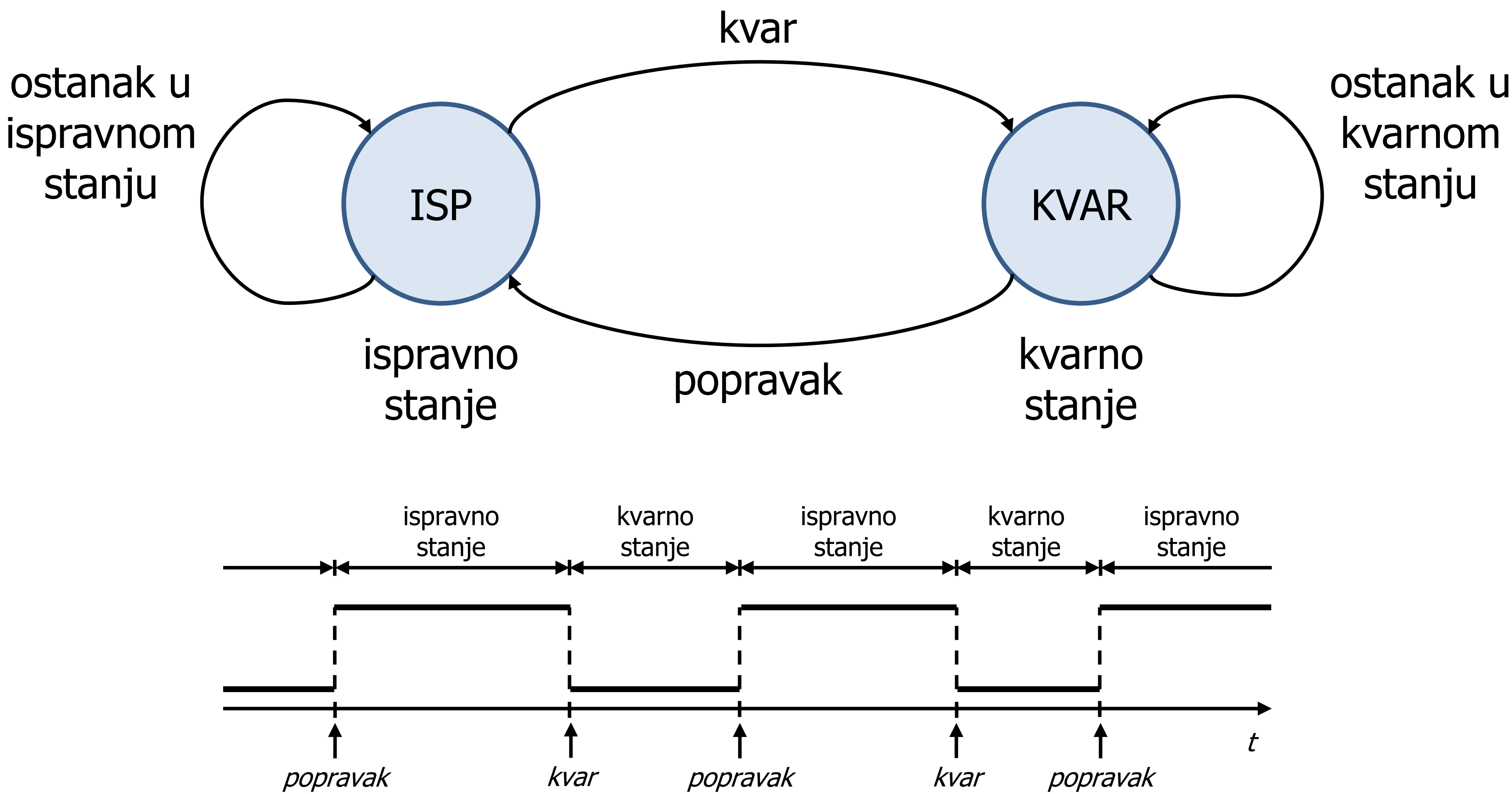


- uporabom više diskova za pohranu jedne datoteke ubrzali smo pristup datoteci
- time smo riješili problem sporih diskova, ali stvorili smo novi problem:
  - uporaba većeg broja diskova povećava vjerojatnost pojave kvarova
- Rješenje: uvesti zalihost i oporavak od kvarova
- zalihost: ↑ povećanje pouzdanosti

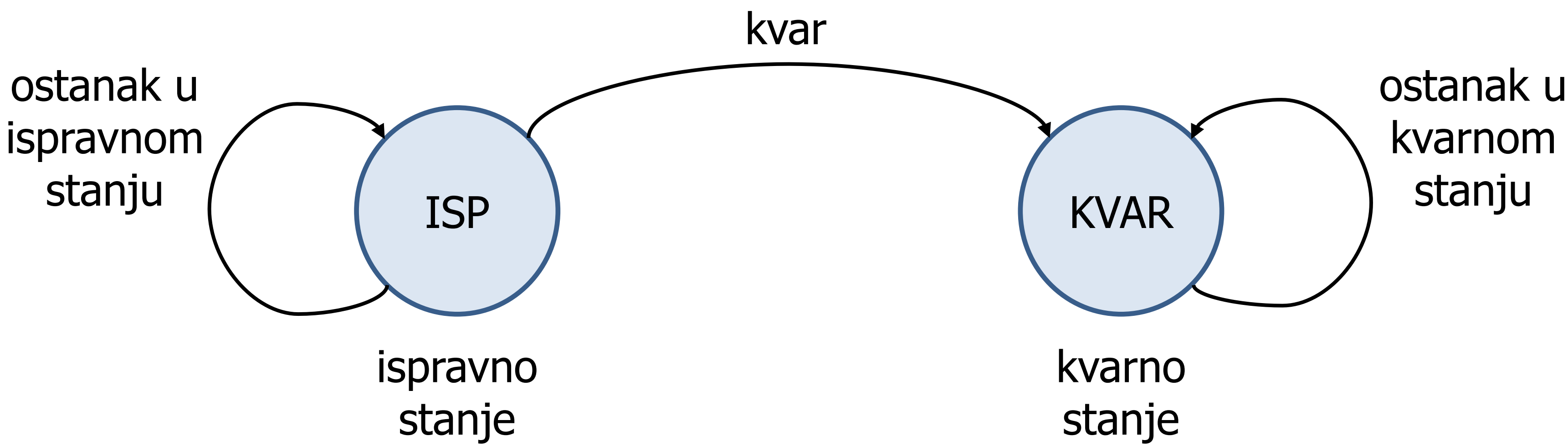
  ↓ smanjenje efektivne brzine prijenosa

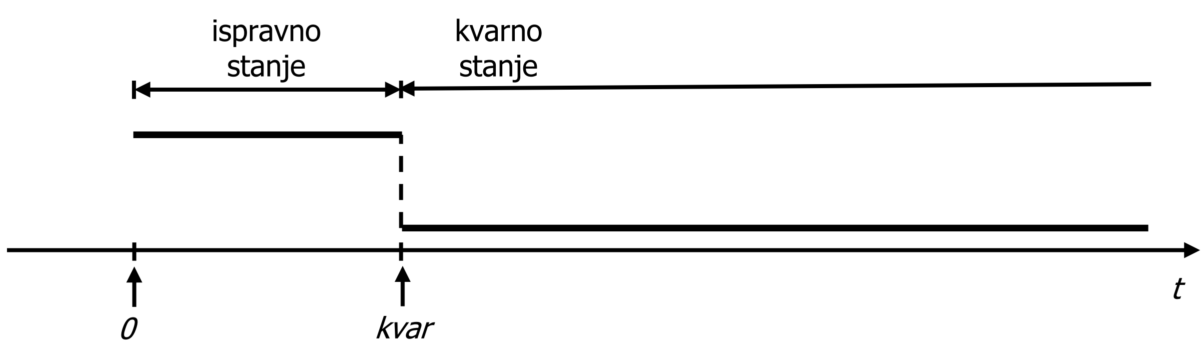
=> postizanje veće brzine i veće pouzdanosti su dva kontradiktorna zahtjeva

Popravljive komponente



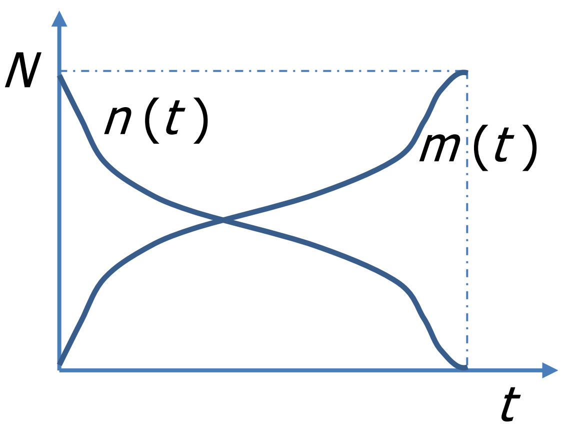
Nepopravljive komponente





Pouzdanost i nepouzdanost nepopravljivih komponenti

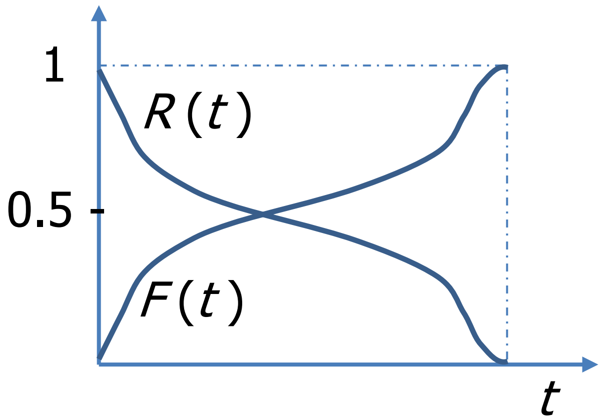
- Pouzdanost (engl. *reliability*) $R(t)$ neke nepopravljive komponente je vjerojatnost da se ona u trenutku t nalazi u ispravnom stanju
  - vjerojatnost da se do trenutka t nije dogodio događaj kvara
  - proporcionalna dijelu populacije koja će ostati u ispravnom stanju do trenutka t
- Nepouzdanost (engl. *unreliability*) $F(t)$ neke nepopravljive komponente je vjerojatnost da se ona u trenutku t nalazi u kvarnom stanju
  - vjerojatnost da se do trenutka t dogodio događaj kvara
  - proporcionalna dijelu populacije koja će prijeći u kvarno stanje do trenutka t



$$n(t) + m(t) = N \quad \Rightarrow \quad \frac{n(t)}{N} + \frac{m(t)}{N} = 1$$

$$\boxed{R(t) + F(t) = 1}$$

Funkcija gustoće vjerojatnosti kvarenja $f(t)$ i prosječno vrijeme do pojave kvara (engl. *Mean Time to Failure* - *MTTF*)

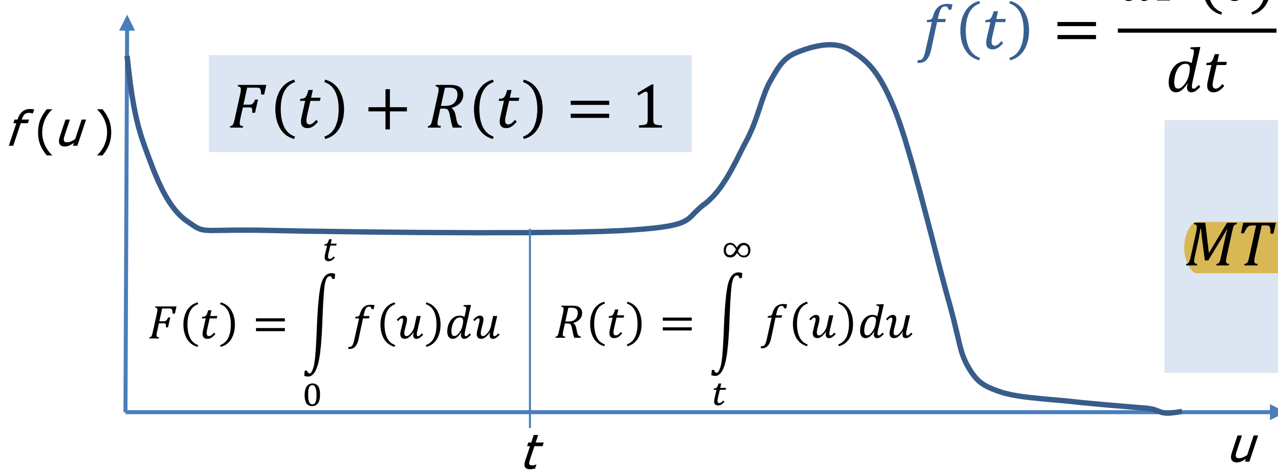


$$\lim_{t\to 0} R(t) = 1 \qquad \lim_{t\to 0} F(t) = 0$$

$$\lim_{t\to \infty} R(t) = 0 \qquad \lim_{t\to \infty} F(t) = 1$$

$$f(t) = \frac{dF(t)}{dt}$$

$$F(t) + R(t) = 1$$



$$F(t) = \int_0^t f(u)du \qquad R(t) = \int_t^\infty f(u)du$$

$$MTTF = \int_0^\infty u f(u) du$$

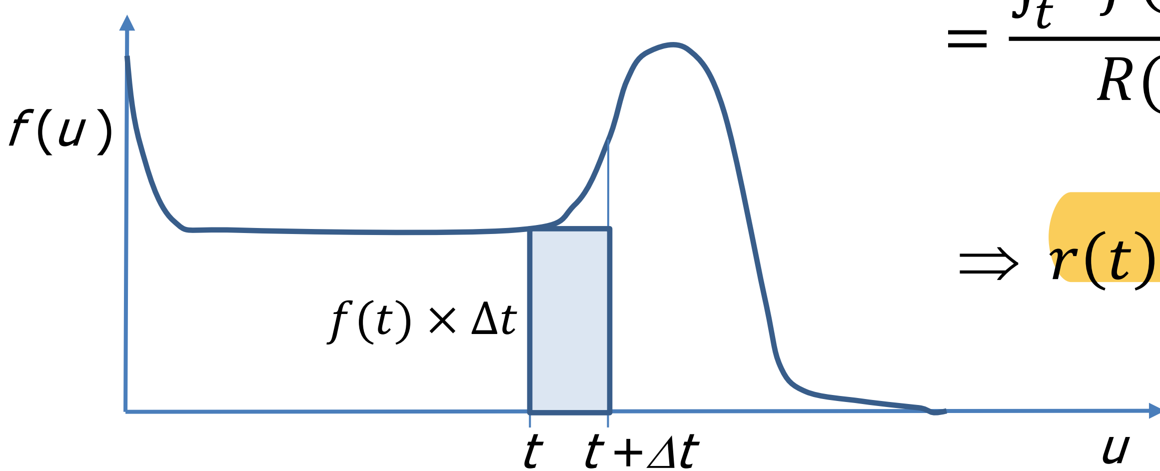
Brzina kvarenja (engl. *failure rate*) $r(t)$

- vjerojatnost pojavljivanja kvara u jedinici vremena u trenutku t i to za onaj dio populacije koji je do tog trenutka preživio

$$r(t) \times \Delta t \approx \frac{m(t+\Delta t) - m(t)}{n(t)} = \frac{dF(t+\Delta t) - F(t)}{R(t)}$$

$$= \frac{\int_t^\infty f(u)du}{R(t)} \approx \frac{f(t) \times \Delta t}{R(t)}$$

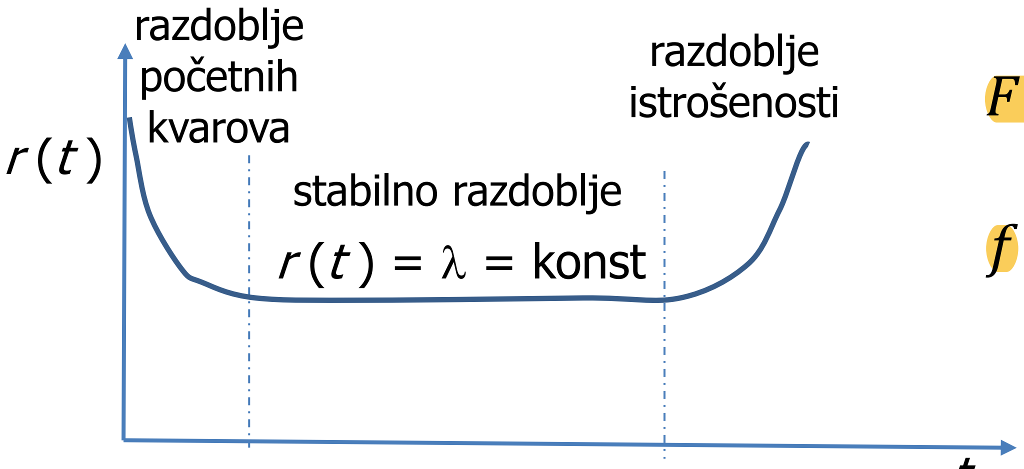
$$\Rightarrow r(t) = \frac{f(t)}{1 - F(t)}$$



$$r(t) = \frac{f(t)}{1 - F(t)} = \frac{\frac{dF(t)}{dt}}{1 - F(t)} = -\frac{d}{dt}\ln[1 - F(t)]$$

$$\Rightarrow \int_0^t r(u)du = -ln[1 - F(t)] \qquad \Rightarrow F(t) = 1 - e^{-\int_0^t r(u)du}$$

U stabilnom razdoblju



$$F(t) = 1 - e^{-\lambda t} \qquad R(t) = e^{-\lambda t}$$

$$f(t) = \lambda \cdot e^{-\lambda t}$$

$$MTTF = \int_0^\infty u\lambda e^{-\lambda t} du = \frac{1}{\lambda}$$

Modeliranje procesa popravljanja komponenti

G (t) - vjerojatnost da je popravak obavljen prije trenutka t

- ima slična svojstva kao funkcija F (t):

$$\lim_{t \to 0} G(t) = 0$$

$$\lim_{t \to \infty} G(t) = 1$$

$$g(t) = \frac{dG(t)}{dt}$$

- funkcija gustoće popravaka

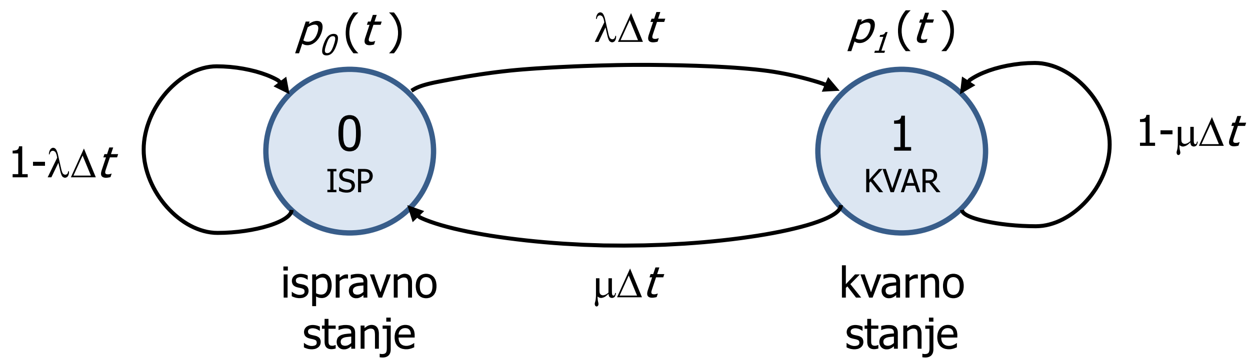
Srednje vrijeme do popravka (engl. *Mean Time to Repair – MTTR*)

$$MTTR = \int_0^\infty u g(u) du$$

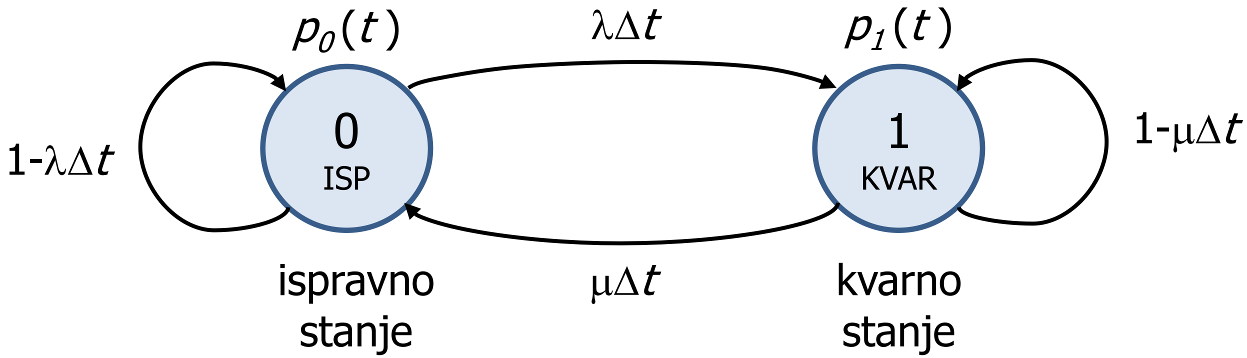
Popravljive i nepopravljive komponente

Svojstvo	Popravak	Kvar
Brzina	$m(t) = \frac{g(t)}{1 - G(t)}$	$r(t) = \frac{f(t)}{1 - F(t)}$
Vjerojatnosti popravka/kvara	$G(t) = 1 - e^{-\int_0^t m(u)du}$	$F(t) = 1 - e^{-\int_0^t r(u)du}$
Konstantna brzina popravljanja μ i kvarenja λ	$G(t) = 1 - e^{-\mu t}$ $g(t) = \mu \cdot e^{-\mu t}$	$F(t) = 1 - e^{-\lambda t}$ $f(t) = \lambda \cdot e^{-\lambda t}$
Prosječno vrijeme trajanja popravka i kvarenja	$MTTR = \frac{1}{\mu}$	$MTTF = \frac{1}{\lambda}$

Model ponašanja popravljive komponente s konstantnim brzinama kvarenja i popravljanja



- $p_0(t)$ – vjerojatnost da je komponenta u trenutku t u ispravnom stanju
- $p_1(t)$ – vjerojatnost da je komponenta u trenutku t u kvaru
- $\lambda\Delta t$ – vjerojatnost da će se u malom intervalu vremena dogoditi kvar
- $\mu\Delta t$ – vjerojatnost da će se u malom intervalu vremena dogoditi popravak



$$p_0(t+\Delta t) = (1-\lambda\Delta t)p_0(t) + \mu\Delta t p_1(t)$$
$$p_1(t+\Delta t) = \lambda\Delta t p_0(t) + (1-\mu\Delta t)p_1(t)$$

$$\frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} = -\lambda p_0(t) + \mu p_1(t)$$

$$\frac{p_1(t+\Delta t) - p_1(t)}{\Delta t} = \lambda p_0(t) - \mu p_1(t)$$

Za $\Delta t \rightarrow 0$ dobiva se sustav diferencijalnih jednadžbi:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t)$$
$$\frac{dp_1(t)}{dt} = \lambda p_0(t) - \mu p_1(t) \qquad [1] \qquad \Rightarrow \begin{bmatrix} \frac{dp_0(t)}{dt} \\ \frac{dp_1(t)}{dt} \end{bmatrix} = \begin{bmatrix} -\lambda & \mu \\ \lambda & -\mu \end{bmatrix} \begin{bmatrix} p_0(t) \\ p_1(t) \end{bmatrix}$$

Opći oblik rješenja sustava diferencijalnih jednadžbi:

$$p_0(t) = C_1 e^{\chi_1 t} + C_2 e^{\chi_2 t}$$
$$p_1(t) = C_3 e^{\chi_1 t} + C_4 e^{\chi_2 t}$$

χ_1 i χ_2 su svojstvene vrijednosti matrice sustava koje se dobivaju iz karakteristične jednadžbe:

$$\left| \chi \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} -\lambda & \mu \\ \lambda & -\mu \end{bmatrix} \right| = \begin{vmatrix} \chi + \lambda & -\mu \\ -\lambda & \chi + \mu \end{vmatrix} = 0, \Rightarrow \chi^2 + (\lambda + \mu)\chi = 0$$

Rješenje kvadratne jednadžbe: $\chi_1 = 0$, $\chi_2 = -(\lambda + \mu)$

Opći oblik rješenja sustava diferencijalnih jednadžbi poprima oblik:

$$p_0(t) = C_1 + C_2 e^{-(\lambda+\mu)t}$$
$$p_1(t) = C_3 + C_4 e^{-(\lambda+\mu)t} \qquad [2]$$

uz početne uvjete: $p_0(0) = 1$, $p_1(0) = 0$

Uvrštavanjem početnih uvjeta u sustav diferencijalnih jednadžbi [1]:

$$\frac{dp_0(0)}{dt} = -\lambda$$
$$\frac{dp_1(0)}{dt} = \lambda$$

deriviranjem [2] i za $t=0$ dobivaju se C_2 i C_4

$$\frac{dp_0(0)}{dt} = -(\lambda+\mu)C_2 e^{-(\lambda+\mu)\cdot 0} = -\lambda \quad \Rightarrow C_2 = \frac{\lambda}{\lambda+\mu}$$
$$\frac{dp_1(0)}{dt} = -(\lambda+\mu)C_4 e^{-(\lambda+\mu)\cdot 0} = \lambda \quad \Rightarrow C_4 = -\frac{\lambda}{\lambda+\mu}$$

Uvrštavanjem početnih uvjeta u [2] dobivaju se preostali koeficijenti C_1 i C_3

$$p_0(0) = C_1 + C_2 = 1 \quad \Rightarrow C_1 + \frac{\lambda}{\lambda+\mu} = 1 \quad \Rightarrow \quad C_1 = \frac{\mu}{\lambda+\mu}$$
$$p_1(0) = C_3 + C_4 \quad \Rightarrow C_3 - \frac{\lambda}{\lambda+\mu} = 0 \quad \Rightarrow \quad C_3 = \frac{\lambda}{\lambda+\mu}$$

Raspoloživost *A(t)* (engl. *availability*) i neraspoloživost *Q(t)* (engl. *unavailability*)

Uvrštavanje dobivenih koeficijenata C_1, C_2, C_3 i C_4 u opći oblik rješenja sustava diferencijalnih dobivaju se izrazi:

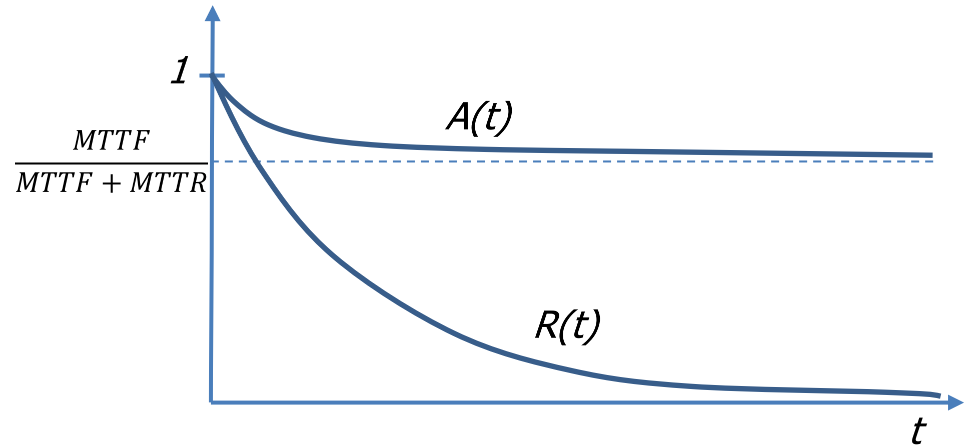
$$p_0(t) = A(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}$$

$$p_1(t) = Q(t) = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} = \frac{\lambda}{\lambda+\mu}\left(1 - e^{-(\lambda+\mu)t}\right)$$

- raspoloživost *A(t)* popravljivih komponenti podsjeća na pouzdanost *R(t)* nepopravljivih komponenti
- neraspoloživost *Q(t)* popravljivih komponenti podsjeća na nepouzdanost *F(t)* nepopravljivih komponenti
- za $\mu = 0$ $A(t) = R(t)$ i $Q(t) = F(t)$

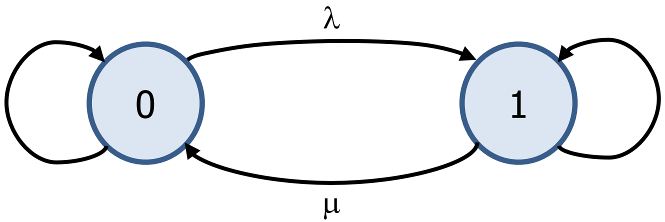
Usporedba raspoloživosti i pouzdanosti

- $\lim_{t \to \infty} R(t) = 0$
- $\lim_{t \to \infty} F(t) = 1$



- $\lim_{t \to \infty} A(t) = \frac{\mu}{\lambda+\mu} = \frac{\frac{1}{MTTR}}{\frac{1}{MTTR}+\frac{1}{MTTF}} = \frac{MTTF}{MTTF+MTTR}$
- $\lim_{t \to \infty} Q(t) = \frac{\lambda}{\lambda+\mu} = \frac{\frac{1}{MTTF}}{\frac{1}{MTTR}+\frac{1}{MTTF}} = \frac{MTTR}{MTTF+MTTR}$

Popravljive komponente



Kvarenje

$r(t) = \lambda$

$R(t) = e^{-\lambda t}$

$F(t) = 1 - e^{-\lambda t}$

$f(u) = \lambda \times e^{-\lambda t}$

$MTTF = \frac{1}{\lambda}$

Popravljanje

$m(t) = \mu$

$G(t) = 1 - e^{-\mu t}$

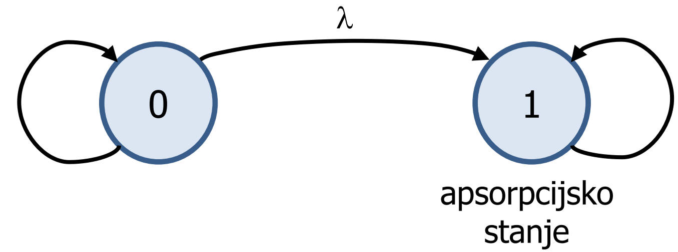
$g(u) = \mu \times e^{-\mu t}$

$MTTF = \frac{1}{\mu}$

$$A(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t}$$

$$Q(t) = \frac{\lambda}{\lambda + \mu}\left(1 - e^{-(\lambda+\mu)t}\right)$$

Nepopravljive komponente



Kvarenje

$r(t) = \lambda$

$R(t) = e^{-\lambda t}$

$F(t) = 1 - e^{-\lambda t}$

$f(u) = \lambda \times e^{-\lambda t}$

$MTTF = \frac{1}{\lambda}$

Popravljanje

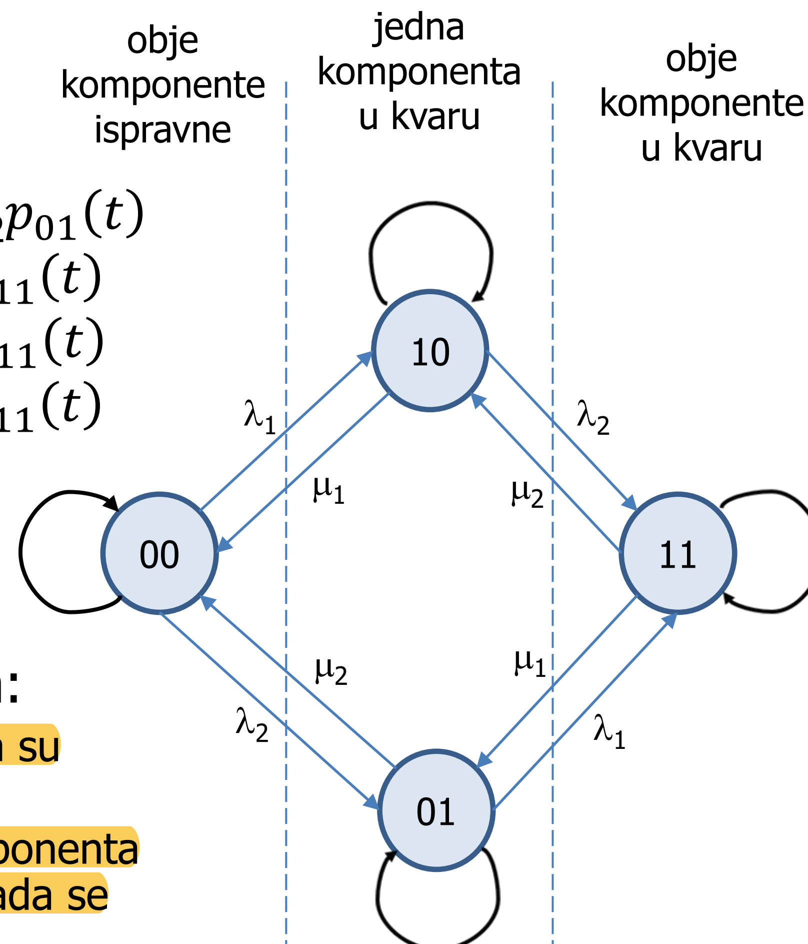
$\mu = 0$

$$A(t) = R(t) = e^{-\lambda t}$$

$$Q(t) = F(t) = 1 - e^{-\lambda t}$$

Modeliranje dvokomponentnog sustava

$$p'_{00}(t) = -(\lambda_1+\lambda_2)p_{00}(t) + \mu_1 p_{10}(t) + \mu_2 p_{01}(t)$$
$$p'_{10}(t) = \lambda_1 p_{00}(t) - (\lambda_2+\mu_1)p_{10}(t) + \mu_2 p_{11}(t)$$
$$p'_{01}(t) = \lambda_2 p_{00}(t) - (\lambda_1+\mu_2)p_{01}(t) + \mu_1 p_{11}(t)$$
$$p'_{11}(t) = \lambda_2 p_{10}(t) + \lambda_1 p_{01}(t) + (\mu_1+\mu_2)p_{11}(t)$$



- neka su sve komponente jednake: $\lambda_1 = \lambda_2 = \lambda$ i $\mu_1 = \mu_2 = \mu$
- promatrat ćemo dva nezavisna slučaja:
  a) sustav je ispravan samo u stanju *00*, tj. kada su obje komponente ispravne
  b) sustav je ispravan i onda kada je jedna komponenta neispravna, tj. sustav se smatra ispravnim kada se nalazi stanjima *00*, *10* ili *01*

a) Sustav je ispravan kada su obje komponente ispravne

$A_S(t) = p_{00}(t)$

$p'_{00}(t) = -(\lambda+\lambda)p_{00}(t)$

za $p_{00}(0) = 1 \quad \Rightarrow \quad p_{00}(t) = e^{-2\lambda t}$

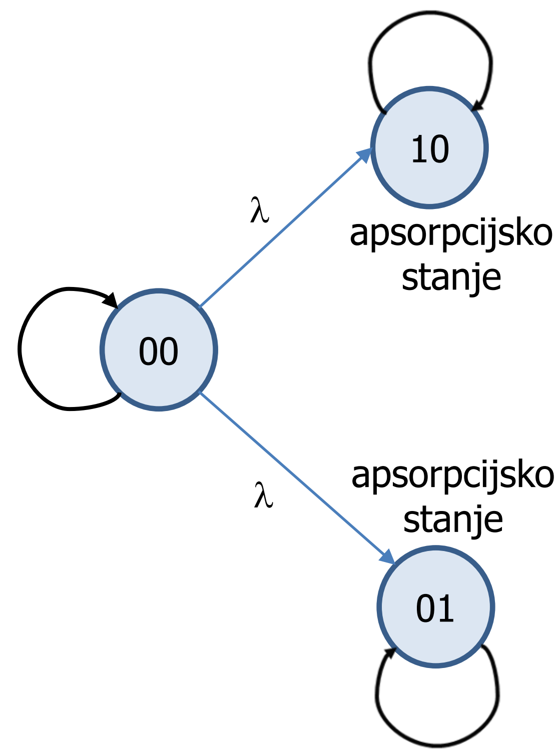
$A_S(t) = e^{-2\lambda t}$

$Q_S(t) = 1 - e^{-2\lambda t}$

$q_S(u) = 2\lambda e^{-2\lambda u}$

$$MTTF_S = \int_0^\infty u 2\lambda e^{-2\lambda u} = \frac{1}{2\lambda}$$

uz $\lambda = \frac{1}{MTTF}$ $\qquad MTTF_S = \frac{1}{2} MTTF$



Sustav s *N* komponenti
a) sustav je ispravan kada su sve komponente ispravne

$$p'_{00}(t) = -(\lambda+\lambda+\cdots+\lambda)p_{00}(t)$$

za $p_{00}(0)=1 \quad \Rightarrow \quad p_{00}(t) = e^{-N\lambda t}$

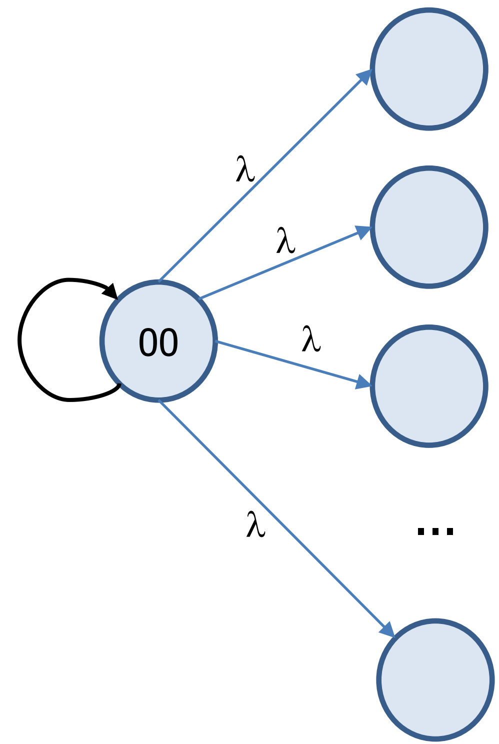
$$A_S(t) = e^{-N\lambda t}$$

$$Q_S(t) = 1 - e^{-N\lambda t}$$

$$q_S(u) = N\lambda e^{-N\lambda u}$$

$$MTTF_S = \int_0^{\infty} uN\lambda e^{-N\lambda u} = \frac{1}{N\lambda}$$

uz $\lambda = \frac{1}{MTTF}$ $\qquad MTTF_S = \frac{1}{N} MTTF$

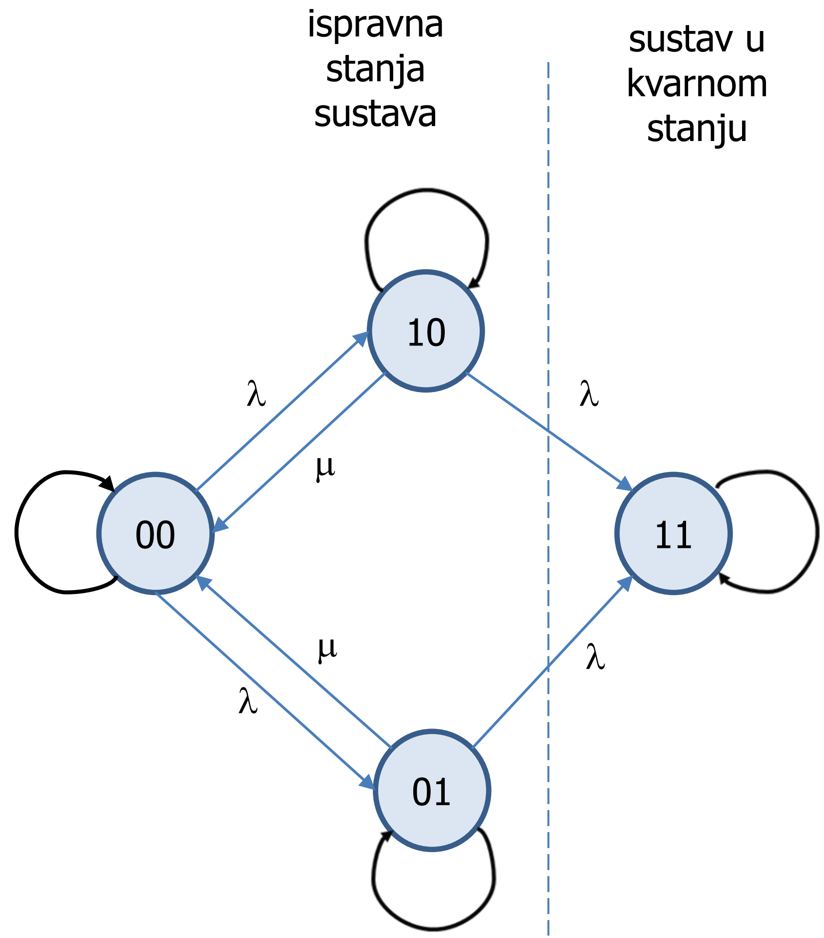


b) Sustav je ispravan i onda kada je jedna komponenta neispravna

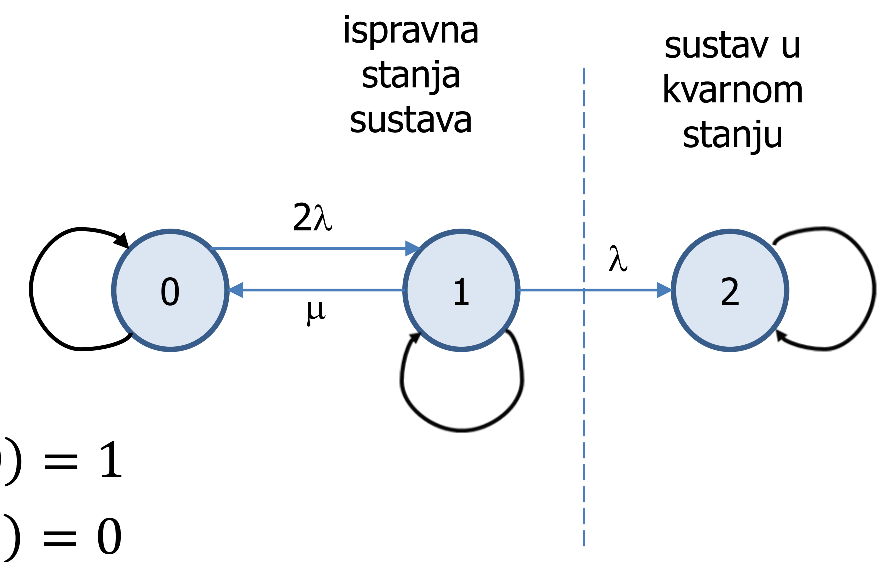
- ispravna stanja su *00, 10* i *10*
- *11* - apsorpcijsko stanje

$$A_S(t) = p_{00}(t) + p_{10}(t) + p_{01}(t)$$
$$Q_S(t) = p_{11}(t)$$

$$p'_{00}(t) = -2\lambda p_{00}(t) + \mu[p_{10}(t) + p_{01}(t)]$$
$$p'_{10}(t) = \lambda p_{00}(t) - (\lambda+\mu)p_{10}(t)$$
$$p'_{01}(t) = \lambda p_{00}(t) - (\lambda+\mu)p_{01}(t)$$
$$p'_{11}(t) = \lambda[p_{10}(t) + p_{01}(t)]$$



Jednostavnije: Markovljev lanac gdje redni broj stanja označava broj neispravnih komponenti



$$A_S(t) = p_0(t) + p_1(t)$$

$$Q_S(t) = p_2(t)$$

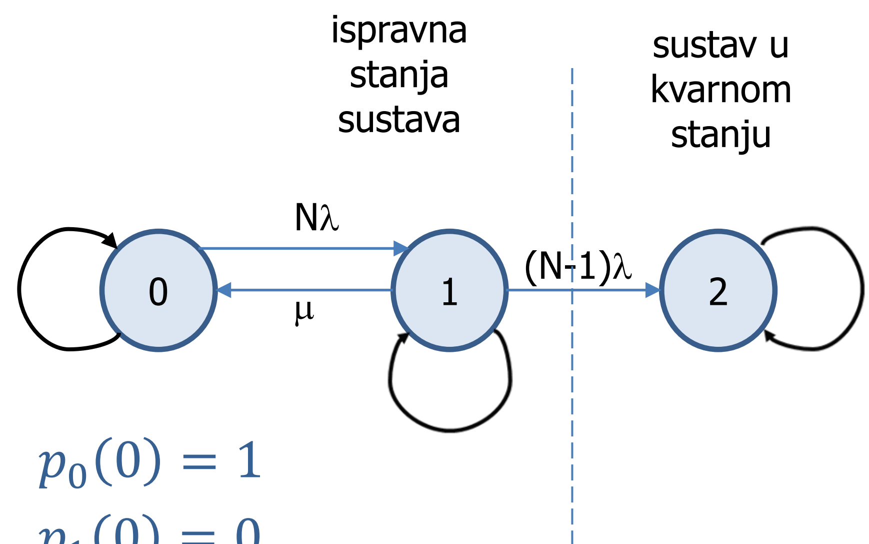
$$p'_0(t) = -2\lambda p_0(t) + \mu p_1(t) \qquad p_0(0) = 1$$

$$p'_1(t) = 2\lambda p_0(t) - (\lambda+\mu) p_1(t) \qquad p_1(0) = 0$$

$$p'_2(t) = \lambda p_1(t) \qquad p_2(0) = 0$$

$$q_S(t) = \frac{dQ_S(t)}{dt} = p'_2(t) = \lambda p_1(t) \qquad MTTF_S = \int_0^\infty u \lambda p_1(u)\, du$$

b) Sustav s *N* komponenti je ispravan i onda kada je jedna komponenta neispravna



$$A_S(t) = p_0(t) + p_1(t)$$
$$Q_S(t) = p_2(t)$$

$$p'_0(t) = -N\lambda p_0(t) + \mu p_1(t) \qquad p_0(0) = 1$$
$$p'_1(t) = N\lambda p_0(t) - [(N-1)\lambda + \mu] p_1(t) \qquad p_1(0) = 0$$
$$p'_2(t) = (N-1)\lambda p_1(t) \qquad p_2(0) = 0$$

$$q_S(t) = \frac{dQ_S(t)}{dt} = p'_2(t) = (N-1)\lambda p_1(t)$$

Traži se $p_1(t)$

Rješenje sustava diferencijalnih jednadžbi svodi se na sustav od dvije jednadžbe s dvije nepoznanice

$$\begin{bmatrix} p'_0(t) \\ p'_1(t) \end{bmatrix} = \begin{bmatrix} -N\lambda & \mu \\ N\lambda & -[(N-1)\lambda + \mu] \end{bmatrix} \begin{bmatrix} p_0(t) \\ p_1(t) \end{bmatrix}$$

Opći oblik rješenja:

$$p_0(t) = A_1 e^{\chi_1 t} + A_2 e^{\chi_2 t}$$

$$p_1(t) = B_1 e^{\chi_1 t} + B_2 e^{\chi_2 t}$$

Izračun svojstvenih vrijednosti χ_1 i χ_1:

$$\begin{vmatrix} \chi + N\lambda & -\mu \\ -N\lambda & \chi + [(N-1)\lambda + \mu] \end{vmatrix} = 0$$

a=1

$$\chi^2 + \underbrace{[(2N-1)\lambda + \mu]}_{b}\chi + \underbrace{N(N-1)\lambda^2}_{c} = 0$$

Vietove formule

$$\chi_1 + \chi_2 = -\frac{b}{a} = -[(2N-1)\lambda + \mu]$$

$$\chi_1 \chi_2 = \frac{c}{a} = N(N-1)\lambda^2$$

Koeficijenti B1 i B2 određuju se iz početnih uvjeta

Iz sustava diferencijalnih jednadžbi poznat je izraz za p'_1 :

$p'_1(t) = \mathrm{N}\lambda p_0(t) - [(N-1)\lambda + \mu]p_1(t)$ uz $p_0(0) = 1$ i $p_1(0) = 0$:

$p'_1(0) = \mathrm{N}\lambda$

a iz općeg rješenja $p_1(t) = B_1 e^{\chi_1 t} + B_2 e^{\chi_2 t}$ dobiva se

$p'_1(t) = \chi_1 B_1 e^{\chi_1 t} + \chi_2 B_2 e^{\chi_2 t}$ odnosno

$$\left.\begin{aligned} p_1(0) &= 0 = B_1 + B_2 \text{ i} \\ p'_1(0) &= \mathrm{N}\lambda = \chi_1 B_1 + \chi_2 B_2 \end{aligned}\right\} \qquad B_1 = \frac{\mathrm{N}\lambda}{\chi_1 - \chi_2} \qquad B_2 = -\frac{\mathrm{N}\lambda}{\chi_1 - \chi_2}$$

Konačno: $$p_1(t) = \frac{\mathrm{N}\lambda}{\chi_1 - \chi_2}\left(e^{\chi_1 t} - e^{\chi_2 t}\right)$$

$$q_S(t) = p'_2(t) = (\mathrm{N}-1)\lambda p_1(t) = \frac{\mathrm{N}(\mathrm{N}-1)\lambda^2}{\chi_1 - \chi_2}\left(e^{\chi_1 t} - e^{\chi_2 t}\right)$$

Prosječno vrijeme do pojave dvostrukog kvara u sustavu

$$MTTF_S = \int_0^\infty u q_s(u) du = \frac{N(N-1)\lambda^2}{\chi_1 - \chi_2}\left[\underbrace{\int_0^\infty u e^{\chi_1 u} du}_{\frac{1}{\chi_1^2}} - \underbrace{\int_0^\infty u e^{\chi_2 u} du}_{\frac{1}{\chi_2^2}}\right]$$

$$MTTF_S = \frac{N(N-1)\lambda^2}{\chi_1 - \chi_2}\left[\frac{1}{\chi_1^2} - \frac{1}{\chi_2^2}\right] = \frac{N(N-1)\lambda^2}{\chi_1 - \chi_2}\left[\frac{\chi_2^2 - \chi_1^2}{\chi_1^2 \chi_2^2}\right] =$$

$$= N(N-1)\lambda^2 \frac{-(\chi_1 + \chi_2)}{(\chi_1 \chi_2)^2} = N(N-1)\lambda^2 \frac{(2N-1)\lambda + \mu}{[N(N-1)\lambda^2]^2} = \frac{(2N-1)\lambda + \mu}{[N(N-1)\lambda^2]} =$$

$$= \frac{1}{N(N-1)}\frac{\mu}{\lambda^2} + \frac{2N-1}{N(N-1)}\frac{1}{\lambda}$$

Prosječno vrijeme do pojave dvostrukog kvara u sustavu

$$MTTF_S == \frac{1}{N(N-1)}\frac{\mu}{\lambda^2} + \frac{2N-1}{N(N-1)}\frac{1}{\lambda}$$

Uvrštavanjem $MTTF = \frac{1}{\lambda}$ i $MTTR = \frac{1}{\mu}$ konačno se dobiva:

$$\boxed{MTTF_S = \frac{1}{N(N-1)}\frac{MTTF^2}{MTTR} + \frac{2N-1}{N(N-1)}MTTF}$$

Primjer 12.1. iz udžbenika

MTTF = 200000 sati = 22.8 godina i

MTTR = 1 sat

Koliko je srednje vrijeme do pojave kvara MTTFS u sustavu s N diskova:

a) ako se kvarom sustava smatra kvar jednog od diskova

b) ako se kvarom sustava smatra kvar dvaju diskova

N	a) $MTTF_S = \frac{1}{N} MTTF$	b) $MTTF_S == \frac{1}{N(N-1)} \frac{MTTF^2}{MTTR} + \frac{2N-1}{N(N-1)} MTTF$
2	100 000 sati	2×10^{10} sati
3	66 666 sati	6.6×10^{9} sati
10	20 000 sati	4.4×10^{8} sati
100	2 000 sati	4 000 000 sati

Načini zalihosne organizacije diskova

RAID - Redundant Array of Independent (Inexpensive) Disks

(*Inexpensive* je iz komercijalnih razloga napušteno.)

RAID 0 — nezalihosna organizacija

- nema zalihosti
- svi diskovi su podatkovni diskovi
- sustav ulazi u kvarno stanje kada se dogodi kvar bilo kojeg diska
- prosječno vrijeme do pojave kvara sustava $MTTF_S$ u višediskovnim se sustavima naziva i prosječnim vremenom do gubitka podataka (engl. *Mean Time to Data Loss – MTTDL*):

$$MTTDL = MTTF_S = \frac{1}{N} MTTF$$

- koristi se dio diska koji je po veličini jednak najmanjem disku u sustavu

Zaključak: – brz, ali se ne smije dogoditi greška

RAID 1 — zrcaljena organizacija

- svaki disk ima svoju kopiju – zrcaljeni disk (engl. *mirrored disk*)
- $MTTDL = \frac{1}{N(N-1)} \frac{MTTF^2}{MTTR} + \frac{2N-1}{N(N-1)} MTTF$ (drugi pribrojnik se zanemaruje)
- svako pisanje se mora obaviti dvaput
- koristi se često za pohranjivanje baza podataka

Zaključak:
– osrednje performanse
– dobra raspoloživost (za čitanje se može koristiti brži disk)
– nedostatak: pola kapaciteta se troši za zalihost

RAID 1 — zanimljivosti

- izraz: $MTTDL = \frac{1}{N(N-1)} \frac{MTTF^2}{MTTR} + \frac{2N-1}{N(N-1)} MTTF$ vrijedi kada sustav ide u kvarno stanje zbog dvostrukog kvara
- primjer slučaja kada sustav RAID 1 ne ide u kvarno stanje unatoč dvostrukom kvaru:

- u sustavu s N diskova dva diska se mogu pokvariti na $\frac{N(N-1)}{2}$ načina od čega samo u $\frac{N}{2}$ slučajeva ide u kvarno stanje kada se pokvare dva diska koja sadrže jednake informacije
- prema tome, izraz za $MTTDL$ treba uvećati za faktor $\frac{\frac{N(N-1)}{2}}{\frac{N}{2}} = N - 1$:

$$MTTDL = \frac{1}{N} \frac{MTTF^2}{MTTR} + \frac{2N-1}{N} MTTF$$

RAID 2 - organizacija zasnovana na Hammingovim kodovima

- za korekciju nakupine od N bitova potrebno je dodati $\log_2 N + 1$ korekcijskih bitova (za N = 4 potrebna 3 korekcijska bita, a za N = 128 potebno ih je 8)
- korekcijski bitovi omogućuju
  - određivanje diska na kojem je nastala pogreška
  - ispravljanje pogrešnog bita

$$MTTDL = \frac{1}{N(N-1)} \frac{MTTF^2}{MTTR}$$

Zaključak: – ne koristi se jer se može na drugi način ustanoviti koji se disk pokvario

RAID 3 - paritetna organizacija sitne zrnatosti

- za korekciju jednostrukih pogrešaka koristi se jedan paritetni disk (jer polažaj pokvarenog diska znamo!)
- pojas ≡ veličina sektora
- pisanje u jedan od pojaseva zahtijeva četiri pristupa do diskova (zbog izračunavanja paritetnog zaštitnog sadržaja):
  1. čitanje starih podataka
  2. čitanje iz paritetnog pojasa
  3. pisanje novih podataka
  4. pisanje u paritetni pojas

Zaključak: – ne koristi se jer paritetni disk može postati usko grlo

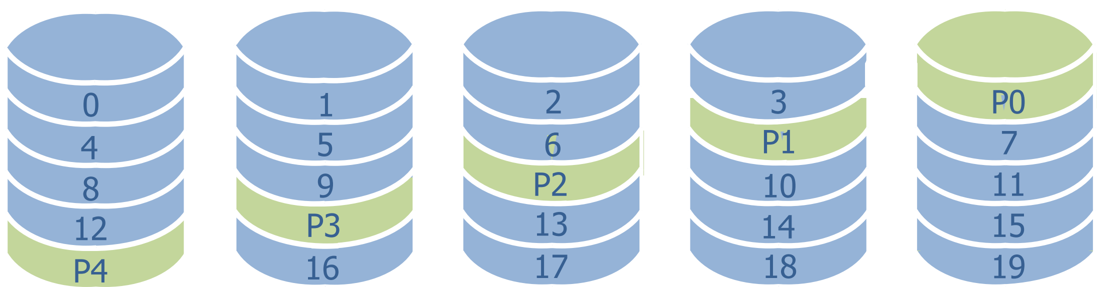
RAID 4 - paritetna organizacija krupne zrnatosti

- kao i kod RAID 3 za korekciju jednostrukih pogrešaka se također koristi jedan paritetni disk, ali su pojasevi veći
- pri kvaru jednog diska mogu se rekonstruirati podaci pokvarenog pojasa uz pomoć pripadnog pojasa paritetnog diska (primjerice, paritetni pojas *P0* štiti podatkovne pojaseve 0, 1, 2 i 3).

Zaključak: – ne koristi se ni RAID 3 ni RAID 4 jer paritetni disk može postati usko grlo

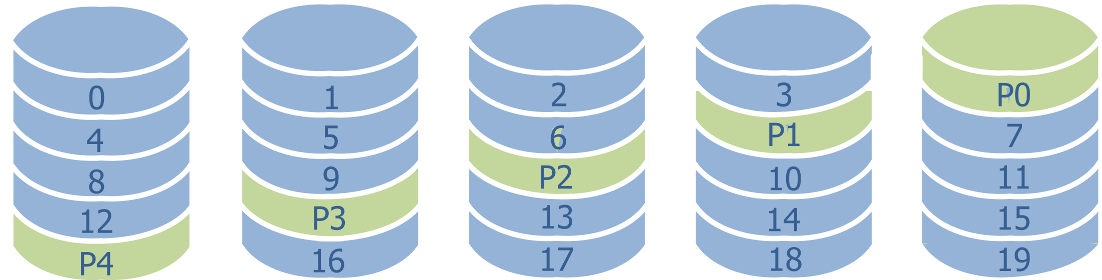
RAID 5 - paritetna organizacija krupne zrnatosti s raspodijeljenim paritetnim pojasevima

- paritetni pojasevi su raspodijeljeni po svim diskovima
- najprikladniji je način razmještaja pojaseva prikazan slikom (zove se *left-asymetric parity distribution*)

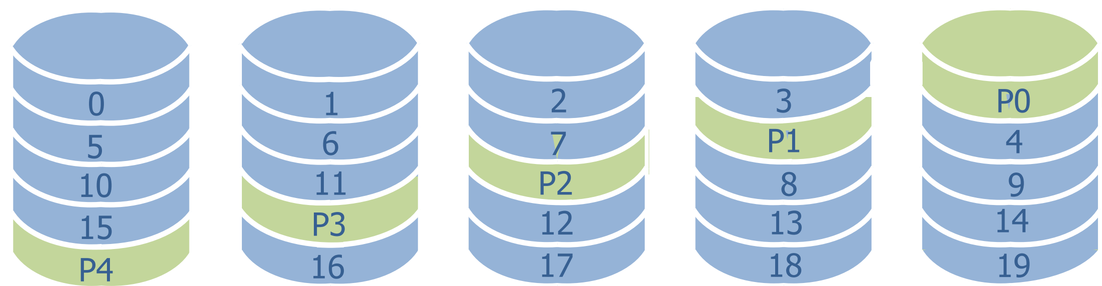


$$MTTDL = \frac{1}{N(N-1)} \frac{MTTF^2}{MTTR}$$

*left-**a**symetric*
parity distribution:



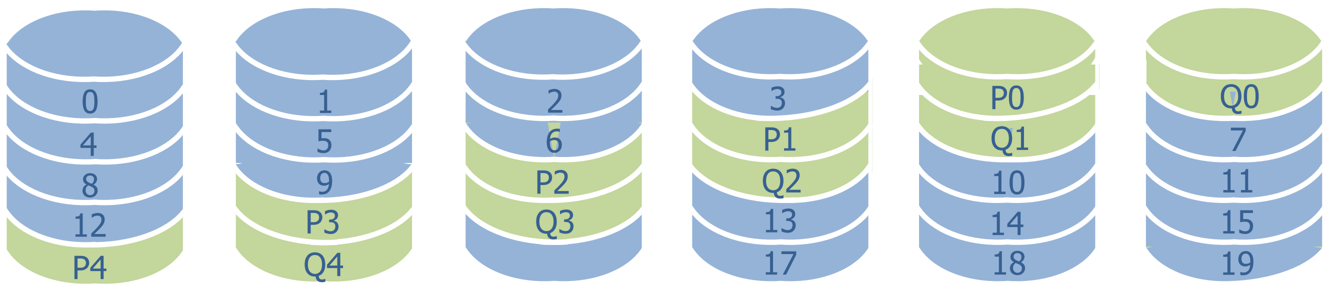
left-symetric
parity distribution:



Zaključak:
- izvrsna raspoloživost
- pisanje sporo kao i kod RAID 4
- RAID 1 i RAID 5 se najčešće upotrebljavaju ali je iskoristivost kapaciteta mnogo veća u RAID 5 organizaciji

RAID 6 – organizacija sa zaštitom od dvostrukog kvara (*P+Q* zalihost)

- podloga za RAID 6 organizaciju su Reed-Solomonovi kodovi
- postoje dva zaštitna diska (odnosno dva zaštitna pojasa za svaku zaštićenu skupinu diskova koji se ravnomjerno raspoređuju po svim diskovima)



$$MTTDL = \frac{1}{N(N-1)(N-2)} \frac{MTTF^3}{MTTR^2}$$

Višerazinski RAID sustavi

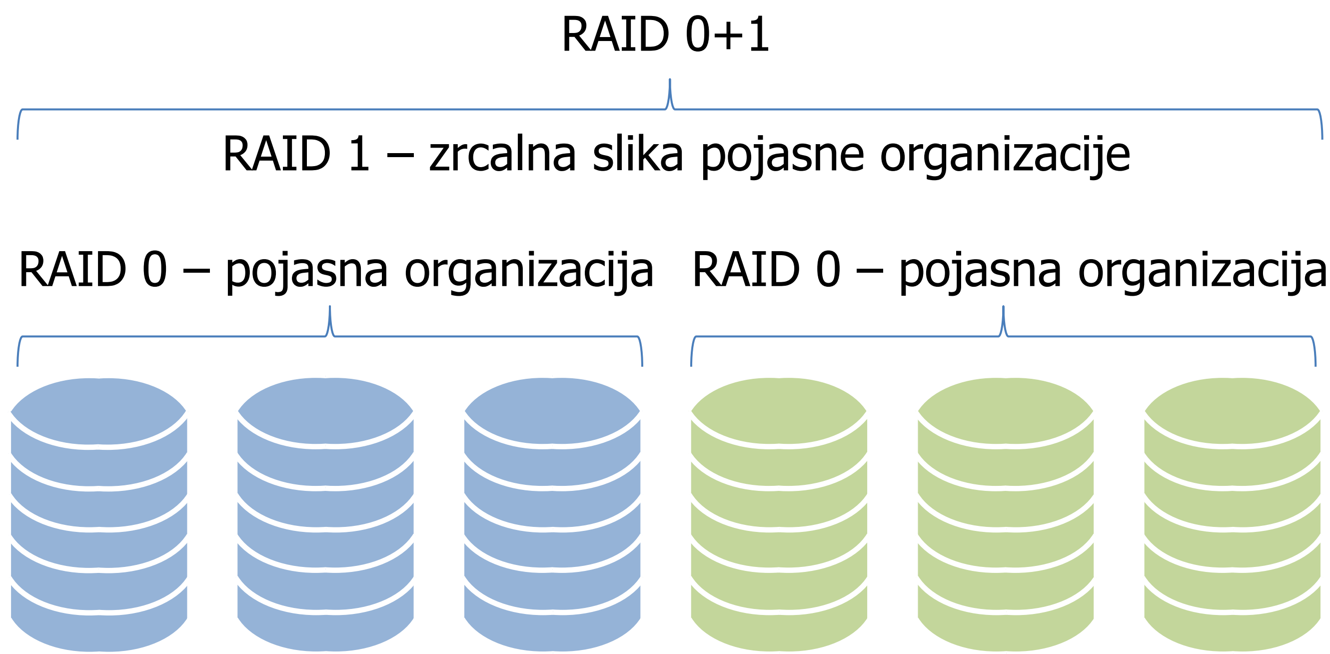
RAID 0+1

RAID 10

RAID 0+1

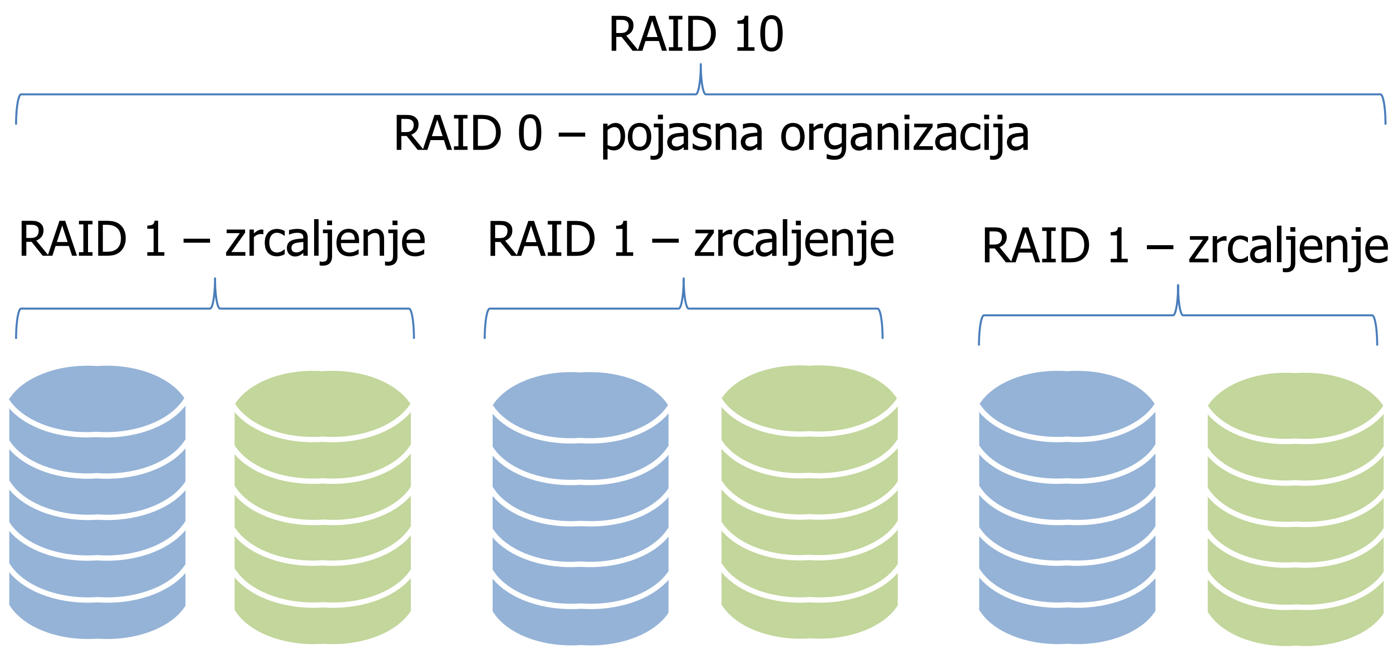
ili RAID 01 ili RAID 0/1

- engl. *mirror of stripes (mirrored stripes)*



RAID 10
ili RAID 1+0 ili RAID 1/0

- engl. *stripe of mirrors (striped mirrors)*

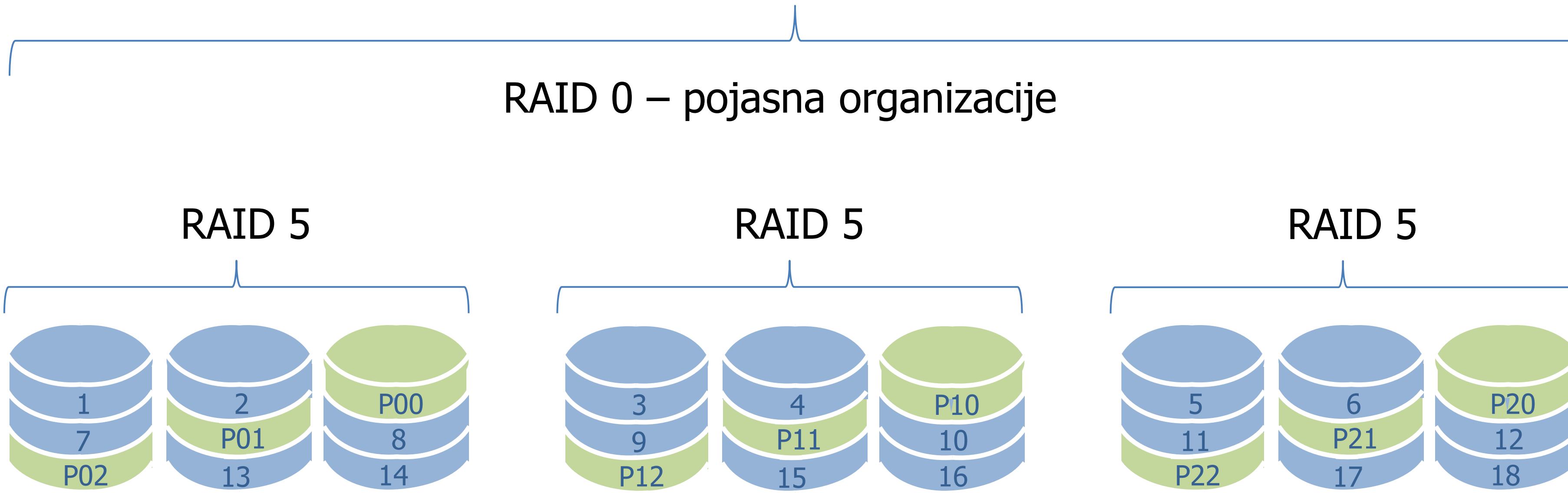


Usporedba sustava RAID 10 i RAID 0+1

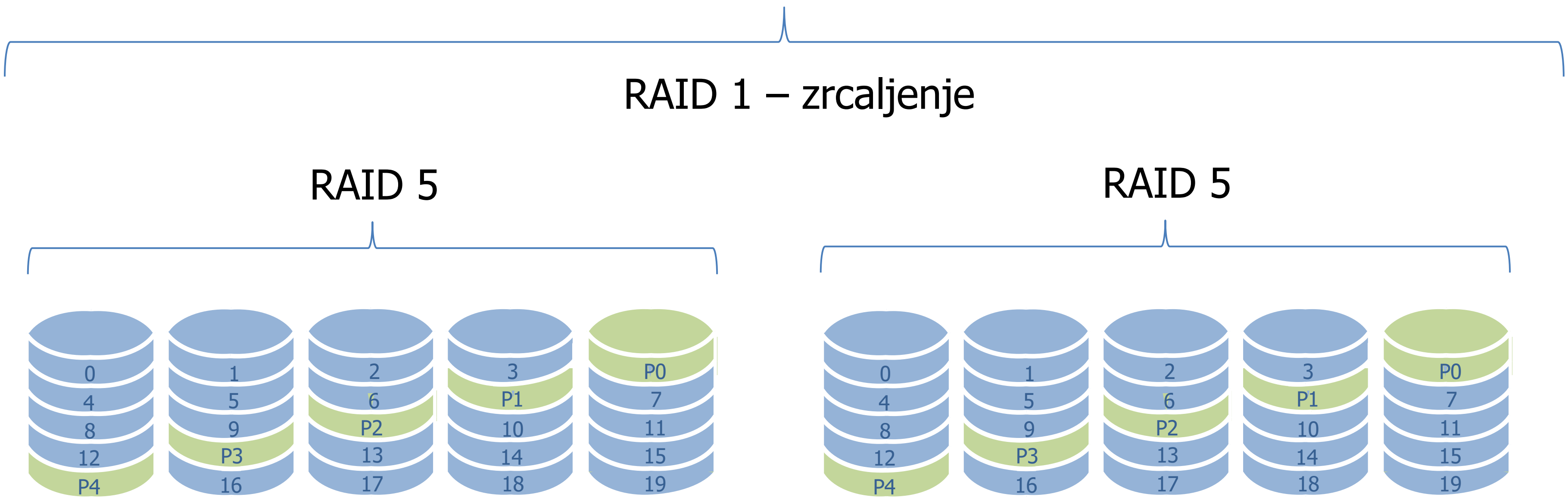
- na prvi pogled izgleda kao da nema razlike
- razlika je u pouzdanosti: RAID 10 je pouzdaniji sustav!
  - u svakoj grupi sustava RAID 10 može se pokvariti 1 disk a da nije došlo do gubljenja podataka jer u svakoj grupi postoji zrcalni disk
    - za 3 grupe po 2 diska ukupno se može pokvariti 3 diska na $2^3=8$ načina po 1 disk iz svake grupe a da sustav ne ide u kvarno stanje
  - u RAID 0+1 sustavu uvijek su dvije grupe diskova (original i zrcalna slika) i do gubitka podataka dolazi ako dođe do kvara obje grupe
    - primjerice, ako dođe do kvara 3 diska u RAID 0+1 sustavu s 2 grupe po 3 diska, podaci ostaju sačuvani u samo 2 slučaja: kada se pokvare svi diskovi iz jedne ili druge grupe

RAID 50

ili RAID 5+0



RAID 51


RAID 1 – zrcaljenje
RAID 5
RAID 5
0
4
8
12
P4
1
5
9
P3
16
2
6
P2
13
17
3
P1
10
14
18
P0
7
11
15
19
0
4
8
12
P4
1
5
9
P3
16
2
6
P2
13
17
3
P1
10
14
18
P0
7
11
15
19

Dodatni slajdovi

Primjer 12.2. iz udžbenika

- neka se neki sustav sastoji od K skupina od G diskova što čini ukupno $N = K \times G$ diskova
- svaka skupina od G diskova zaštićena je RAID 5 načinom u kojem očekivano vrijeme do gubitka podataka iznosi:

$$MTTDL_G = \frac{1}{N(N-1)} \frac{MTTF^2}{MTTR}$$

- ako se pojavi kvar jedne od K grupa (što znači dvostruki kvar unutar te grupe), onda imamo gubitak podataka te je

$$MTTDL_N = \frac{1}{K} \frac{1}{G(G-1)} \frac{MTTF^2}{MTTR} = \frac{1}{N(G-1)} \frac{MTTF^2}{MTTR}$$

Primjer 12.2. iz udžbenika proširen sa sustavom RAID 6

- primjerice uz parametre sustava:
  MTTF = 200 000 sati
  MTTR = 1 sat
  N = 96
  G = 16

dobiva se za RAID 5 sustav:

$$MTTDL_{96} = \frac{1}{96 \times (16-1)} \frac{(2 \times 10^5)^2}{1}$$

$$= 2.78 \times 10^7\ sati$$

$$= 3170\ godina$$

- uz jednake parametre, ali uz primjenu RAID 6 organizacije sustava prosječno vrijeme do pojave gubitka podataka iznosi:

$$MTTDL_N = \frac{1}{N(G-1)(G-2)} \frac{MTTF^3}{MTTR^2}$$

odnosno

$$MTTDL_{96} = \frac{1}{96 \times (16-1)(16-2)} \frac{(2 \times 10^5)^3}{1}$$

$$= 3.97 \times 10^{11}\ sati$$

$$= 45\,300\,000\ godina$$

Korisni kapacitet RAID sustava

Neka se RAID sustav sastoji od N

- istovrsnih diskova kapaciteta C ili
- različitih diskova gdje je disk s najmanjim kapacitetom C_{MIN} pa je $C = C_{MIN}$

RAID 0	RAID 1	RAID 5	RAID 6
$N \times C$	$\frac{N}{2} \times C$	$(N-1) \times C$	$(N-2) \times C$